

Crystal Structure of Lactose Synthase Reveals a Large Conformational Change in its Catalytic Component, the β 1,4-Galactosyltransferase-I

B. Ramakrishnan^{1,2} and Pradman K. Qasba^{1*}

¹Structural Glycobiology Section

²Intramural Research Support Program-SAIC, Laboratory of Experimental and Computational Biology, CCR NCI, Frederick MD 21702, USA

The lactose synthase (LS) enzyme is a 1:1 complex of a catalytic component, β 1,4-galactosyltransferase (β 4Gal-T1) and a regulatory component, α -lactalbumin (LA), a mammary gland-specific protein. LA promotes the binding of glucose (Glc) to β 4Gal-T1, thereby altering its sugar acceptor specificity from *N*-acetylglucosamine (GlcNAc) to glucose, which enables LS to synthesize lactose, the major carbohydrate component of milk. The crystal structures of LS bound with various substrates were solved at 2 Å resolution. These structures reveal that upon substrate binding to β 4Gal-T1, a large conformational change occurs in the region comprising residues 345 to 365. This repositions His347 in such a way that it can participate in the coordination of a metal ion, and creates a sugar and LA-binding site. At the sugar-acceptor binding site, a hydrophobic *N*-acetyl group-binding pocket is found, formed by residues Arg359, Phe360 and Ile363. In the Glc-bound structure, this hydrophobic pocket is absent. For the binding of Glc to LS, a reorientation of the Arg359 side-chain occurs, which blocks the hydrophobic pocket and maximizes the interactions with the Glc molecule. Thus, the role of LA is to hold Glc by hydrogen bonding with the O-1 hydroxyl group in the acceptor-binding site on β 4Gal-T1, while the *N*-acetyl group-binding pocket in β 4Gal-T1 adjusts to maximize the interactions with the Glc molecule. This study provides details of a structural basis for the partially ordered kinetic mechanism proposed for lactose synthase.

Keywords: lactose synthase; crystal structure; substrate binding; conformational changes; modulation

*Corresponding author

Introduction

Lactose synthase (LS; EC 2.4.1.22) consists of two components: β 1,4-galactosyltransferase (β 4Gal-T1; EC 2.4.1.90/38) is one in which the catalytic activity resides; the other, α -lactalbumin (LA), modulates the activity of the former. β 4Gal-T1 is a widely distributed, Golgi resident, type-II membrane protein (~45 kDa).¹ LA, on the other hand, is a mammary gland-specific, Ca²⁺-binding protein (~14 kDa) that is expressed, and in large quantities, only during lactation.² In the presence of Mn²⁺, β 4Gal-T1 catalyzes the transfer of the galactosyl residue from UDP-galactose (UDP-Gal) to *N*-acetyl-glucosa-

mine (GlcNAc) in either its free form or linked to an oligosaccharide. LA modifies Gal acceptor specificity of β 4Gal-T1 such that it transfers Gal to glucose as well, to synthesize lactose, the "milk sugar". In the absence of LA, β 4Gal-T1 has low affinity for Glc, with a K_m of ~2 M. In the presence of LA, however, the K_m for Glc is reduced by 1000-fold, so that its physiological concentrations are adequate for LS activity.³

LA has long been known to share a similarity with lysozyme at both the genetic and protein structure levels.⁴ Even though these two proteins exhibit close structural similarities, functionally they are quite different. Lysozyme binds and hydrolyzes the β 1-4 glycosidic bond in an oligosaccharide, whereas LA does not even bind the oligosaccharide,⁴ let alone show a catalytic activity. Furthermore, while LA binds to β 4Gal-T1 and modulates its sugar acceptor specificity, lysozyme shows no affinity for β 4Gal-T1.

Abbreviations used: LS, lactose synthase; LA, α -lactalbumin; β 4Gal-T1, β 1,4-galactosyltransferase.

E-mail address of the corresponding author: qasba@helix.nih.gov

In humans, β 1,4-galactosyltransferase (β 4Gal-T) consists of a family of seven enzymes, β 4Gal-T1 to β 4Gal-T7, that have high protein sequence homology.^{5,6} These enzymes are expressed in different tissues and show differences in the acceptor specificity. In the presence of Mn^{2+} , all these enzymes transfer Gal from UDP-Gal to different sugar acceptors, generating a β 1-4 linkage. During lactation, only β 4Gal-T1 is expressed in the mammary gland. Besides β 4Gal-T1, some other members of this enzyme family do interact with LA but they exhibit properties that are different. For example, β 4Gal-T4 does interact with LA but it synthesizes *N*-acetyl-lactosamine effectively but not lactose.⁷ β 4Gal-T6, which transfers Gal to the Glc residue of glucosylceramide to synthesize lactosylceramide,⁸ does not interact with LA.

Although the crystal structures of both LA⁹ and the catalytic domain of free β 4Gal-T1¹⁰ have been known for some time, their mode of interaction has not been elucidated. Neither the sugar-acceptor site nor the LA-binding site was clearly identified in the crystal structure of β 4Gal-T1.¹⁰ For a better knowledge of the interactions between these proteins, which was deemed essential for the understanding of the mechanism of action of lactose synthase, we undertook the crystal structure determination of LS. We solved the LS crystal structure by MAD methods, using Se-methionine contained in the catalytic domain of β 4Gal-T1 (~33 kDa) complexed with LA. We studied the crystal structure of LS and those of its ternary complexes with its preferred substrates. These studies revealed a large conformational change in the region comprising residues 345 to 365 of β 4Gal-T1 when compared to the crystal structure of β 4Gal-T1.¹⁰ This conformational change creates the sugar-acceptor and LA-binding sites, and redefines the catalytic pocket.

Results and Discussions

Overall structure of LA and β 4Gal-T1 molecules

Table 1 lists single crystal parameters and final refinement statistics on the lactose synthase with various substrates. Two LA and two β 4Gal-T1 molecules are present in the asymmetric unit and they are related by a pseudo 2-fold symmetry. Each LA molecule is bound to one β 4Gal-T1 molecule with either one monosaccharide or uridine diphosphate (Figure 1 shows LS with GlcNAc), forming a 1:1 protein complex of LS.

The overall structure of the catalytic domain of the β 4Gal-T1 in LS is similar to the reported structure for free β 4Gal-T1,¹⁰ except for the region comprising residues 345 to 365 (Figure 2). In order to distinguish the two conformations, we refer to the previously reported structure of free β 4Gal-T1 as conformation I and the structure of β 4Gal-T1 found in LS as conformation II. The mean r.m.s. deviation for the corresponding C^α atoms in these two structures in this flexible region, residues 345 to 365, of these two structures is 9.8 Å with a maximum of 20 Å for Lys352. In contrast, the value for the rest of the molecule is only 0.6 Å. This region in conformation I exists mainly as a loop, whereas in conformation II residues 345 to 358 form a loop, while the remaining residues, 359 to 365, form an α -helix. These changes cause alterations in the structure of the catalytic pocket. In conformation I, this has been defined as a deep open pocket, 13 Å in diameter.¹⁰ In contrast, in conformation II it changes to an extended cleft. Other significant conformational difference involves Gly313 and Trp314. The side-chain of Trp314, which was outside the pocket is inside the pocket in conformation II (Figure 2). These observations are consistent with the inference drawn earlier regarding the presence of a Trp residue at the active site.¹¹ UV inactivation of the enzyme caused by the destruction of a single

Table 1. Single crystal parameters and final refinement statistics

| Crystals grown in the presence of | (Se-Met)-LS·UDP-Gal·Mn ²⁺ ·4F-Glc | LS·GlcNAc | LS·Glc | LS·UDPGal·Mn ²⁺ |
|-----------------------------------|--|-------------|-------------|----------------------------|
| Substrate found | UDP | GlcNAc | Glc | UDP·Mn ²⁺ |
| Unit cell parameters | | | | |
| <i>a</i> (Å) | 55.2 | 57.2 | 57.2 | 55.5 |
| <i>b</i> (Å) | 98.9 | 96.4 | 93.9 | 99.2 |
| <i>c</i> (Å) | 102.3 | 99.5 | 100.1 | 102.3 |
| β (deg.) | 103.9 | 101.0 | 101.5 | 104.1 |
| Resolution (Å) | 2.5 | 2.0 | 2.0 | 2.0 |
| Reflections collected | ~110,000 | 153,624 | 119,864 | 224,308 |
| Unique reflections (%) | 37,099 (98) | 68,631 (94) | 57,639 (81) | 71,688 (99) |
| R_{sym} | 0.050 | 0.070 | 0.089 | 0.082 |
| Reflections used | 33,047 | 62,687 | 51,690 | 54,680 |
| Final R factor/ R_{free} | 0.25/0.30 | 0.25/0.29 | 0.24/0.27 | 0.26/0.29 |
| r.m.s. deviations on | | | | |
| Bonds (Å) | 0.008 | 0.007 | 0.006 | 0.006 |
| Angles (deg.) | 1.5 | 1.4 | 1.4 | 1.5 |

The crystals belong to a monoclinic system, with space group $P2_1$. The structure of (Se-Met)-LS·UDP-Gal·Mn²⁺·4F-Glc was determined at 2.5 Å resolution using MAD methods. The structures of LS complexes with various substrates were then determined at 2.0 Å resolutions. LS is a 1:1 complex between bovine β 4Gal-T1, residues 130 to 402, and mouse LA. 4F-Glc is 4fluoro-glucose.

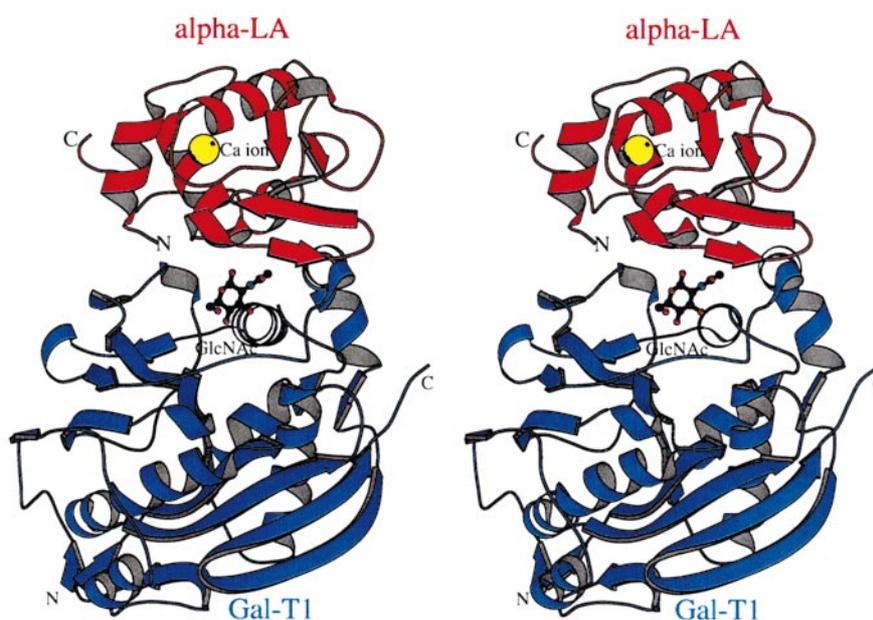


Figure 1. The molecular structure of LS, a 1:1 complex between the catalytic domain of bovine $\beta 4\text{Gal-T1}$ (residue 130 to 402)²⁶ and mouse LA. The complex is shown with the acceptor GlcNAc. The interaction of LA with $\beta 4\text{Gal-T1}$ is mostly near the acceptor site, away from both the N and C termini of $\beta 4\text{Gal-T1}$.

Trp residues was prevented when UDP-Gal and Mn^{2+} were present.¹¹ In order to exclude the possibility that the *in vitro* folding used in the current investigations caused the conformational changes, in a separate study we determined the structure of $\beta 4\text{Gal-T1}$, purified from NSO-GT cell lines obtained from ATCC,¹⁰ and observed the same structure in LS-complex as that described here.

Although the crystal structure of mouse LA *per se* has not been determined, its structure in the LS complex in the present study is very similar to the published structure of free LA from other species.⁹ LA has a flexible region, residues 105-111, which is

known to be important for $\beta 4\text{Gal-T1}$ binding,¹² and adopts a loop or helix conformation, depending upon the pH of the crystallization medium.¹³ The present study shows that this region of mouse LA adopts a helix conformation when it is bound in LS (Figure 1).

Molecular interactions between LA and $\beta 4\text{Gal-T1}$ molecules

The interactions between LA and $\beta 4\text{Gal-T1}$ molecules are primarily hydrophobic (Figure 3). The distances between the possible interacting atoms in the two protein molecules are listed in Table 2. The

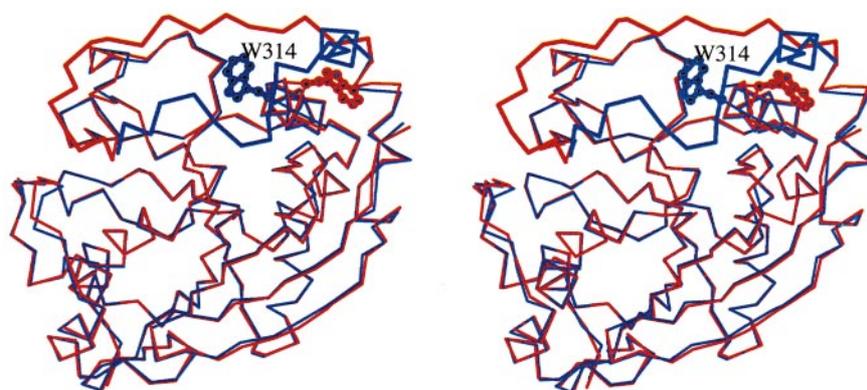


Figure 2. Superposition of C^α atoms of the $\beta 4\text{Gal-T1}$ structure from the LS (blue; present study) with conformation I (see the text) of $\beta 4\text{Gal-T1}$ ¹⁰ (red). The mean r.m.s. deviation between the two proteins is only 0.6 Å, except for the segment comprising residues 345 to 365. In this segment, the r.m.s. deviation is 9.8 Å with a maximum of 21 Å. It lies on the opposite side of the catalytic pocket and within this region residues 359 to 365 form an α -helix in LS (blue). Secondly, the orientation of the Trp314 side-chain in $\beta 4\text{Gal-T1}$ in LS complex conformation II (blue) is facing towards inside the catalytic pocket, whereas it is towards the outside in conformation I (red).

Table 2. Interatomic distances between the residues from LA and β 4Gal-T molecules in the LSCGlcNAc complex interactions with distances less than 3.8 Å are listed

| LA | Atom | β 4Gal-T1 | Atom | Distance (Å) |
|--------|-------------------------------------|-----------------|-------------------------------------|--------------|
| Glu2 | O ^{ϵ1} | Arg346 | N ⁿ | 2.96 |
| Glu2 | O ^{ϵ2} | Arg346 | N ⁿ¹ | 3.55 |
| Phe31 | C ^{ϵ1} | Pro285 | C ^{β} | 3.65 |
| Phe31 | C ^{ϵ2} | Tyr286 | C ^{δ2} | 3.62 |
| | | Tyr286 | C ^{ϵ2} | 3.59 |
| Phe31 | C ^{ζ} | Pro285 | C ^{β} | 3.58 |
| His32 | C ^{ϵ1} | Tyr286 | C ^{δ2} | 3.74 |
| | | Tyr286 | C ^{ϵ2} | 3.64 |
| | | Tyr286 | C ^{ζ} | 3.69 |
| His32 | O | Phe360 | C ^{ζ} | 3.63 |
| Val42 | C ^{γ1} | Glu354 | O ^{ϵ2} | 3.35 |
| Asp44 | O | Glu354 | O ^{ϵ1} | 3.71 |
| Asp44 | O ^{δ1} | Glu354 | O ^{ϵ2} | 3.72 |
| Lys105 | C | Phe360 | C ^{γ} | 3.79 |
| Lys105 | O | Phe360 | C ^{β} | 3.49 |
| | | Phe360 | C ^{δ2} | 3.67 |
| Ala106 | C ^{α} | Phe360 | C ^{δ2} | 3.75 |
| | | Phe360 | C ^{ϵ2} | 3.60 |
| Lys108 | N ^{ζ} | Ala364 | C ^{β} | 3.27 |
| Pro109 | C ^{β} | ILE363 | C ^{γ2} | 3.53 |
| Met110 | C ^{ϵ} | Tyr286 | C ^{ϵ1} | 3.71 |
| | | Asp319 | O ^{δ1} | 3.76 |
| Lys114 | N ^{ζ} | Val287 | C ^{γ1} | 3.78 |
| | | Tyr322 | OH | 3.48 |
| Gln117 | O ^{ϵ1} | Tyr286 | C ^{α2} | 3.46 |
| | | Tyr286 | C | 3.62 |
| | | Val287 | N | 2.93 |
| | | Val287 | C ^{β} | 3.70 |
| Gln117 | N ^{ϵ2} | Val287 | C ^{γ1} | 3.80 |
| | | Gln288 | O ^{ϵ1} | 3.13 |
| Trp118 | C ^{α2} | Tyr286 | C ^{β} | 3.65 |
| Trp118 | C ^{ζ} | Pro285 | O | 3.66 |

hydrophobic patch formed by residues Phe31, His32, Met110, Gln117 and Trp118 in the LA molecule^{12,14} interacts with a corresponding hydrophobic patch in β 4Gal-T1 formed by residues Phe280, Tyr286, Gln288, Tyr289, Phe360 and Ile363 (Figure 3). The D-helix of LA (residues 105 to 111),

interacts with the α 6 helix of β 4Gal-T1 (residues 359 to 365), the Phe360 side-chain of β 4Gal-T1 interacts with the backbone of the D-helix of LA; and the Pro109 side-chain of LA interacts with the backbone of the α 6 helix of β 4Gal-T1. The LA molecule interacts only with β 4Gal-T1 and the acceptor

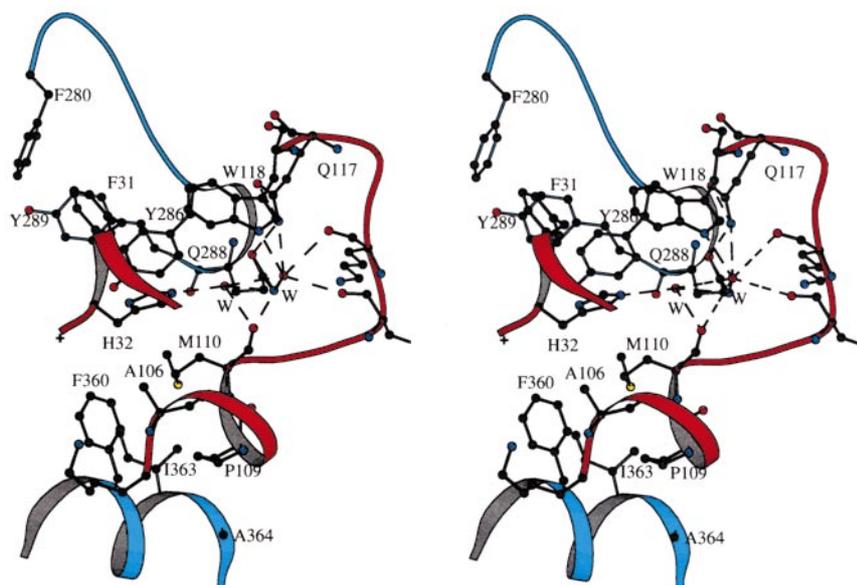


Figure 3. Stereo view of the molecular interactions between β 4Gal-T1 and LA. The secondary structure elements of LA are colored red, while those of β 4Gal-T1 are in blue. The known hydrophobic patch present in LA interacts with the proposed hydrophobic patch present in β 4Gal-T1. The interactions between the two proteins are mostly hydrophobic. Two structural water molecules are trapped between the two proteins, and one of them has extensive hydrogen bonding interactions with the protein molecules.

sugar molecule, and not with the donor substrate, UDP-Gal (see below). These observations substantiate the results of site-directed mutagenesis studies of LA that Ala106, His107, Leu110, Gln117 and Trp118 are important for binding to β 4Gal-T1, and Phe31 and His32 are important for Glc binding.¹⁴ Also, in the LS complex, a total of 1310 Å² solvent-accessible surface area is buried between the two protein molecules. This corresponds to 20% and 11% of the individual solvent-accessible area of LA and β 4Gal-T1 molecules, respectively.

Significance of the two conformational states of β 4Gal-T1 for LA-binding

Cross-linking experiments¹⁵ have established that β 4Gal-T1 binds to LA only in the presence of either sugar acceptors or UDP-GalMn²⁺ (Figure 4(a)). This strongly suggests that in the absence of these ligands, β 4Gal-T1 exists in a form that cannot interact with LA. The crystal structure of free β 4Gal-T1¹⁰ indicates that the flexible region that constitutes residues 345-365 buries the LA-binding site. It appears likely that this state represents conformation I of β 4Gal-T1 (Figure 4(b)). As β 4Gal-T1 interacts with LA only in the presence

of its substrate and because all of our LS crystals could be grown only in the presence of at least one substrate, we conclude that conformation II is the catalytically active conformation in LS (Figure 4(c)). A conformational transition of the flexible region from conformation I to conformation II could expose the LA-binding region on β 4Gal-T1. As suggested by the chemical cross-linking experiments (Figure 4(a)), the observed flexible region conformation in the present structure is likely to be due to the substrate binding and not due to LA-binding. In fact, in the absence of LA, in the crystal structure of β 4Gal-T1 complex with UDP-Gal and Mn²⁺ alone, β 4Gal-T1 has a conformation II similar to that observed in the present LS complex (unpublished results). This conclusion is supported by spectroscopic studies, which suggested that upon UDP-Gal and Mn²⁺ binding, β 4Gal-T1 undergoes significant conformational changes.¹⁶ Although there is no evidence showing conformational changes upon the sugar acceptor binding to β 4Gal-T1, cross-linking of LA and β 4Gal-T1 in the presence of GlcNAc (Figure 4(a)) suggests that β 4Gal-T1 most likely undergoes conformational changes upon GlcNAc binding. Since the conformation of the flexible region in the crystal structure of LS·UDP·Mn²⁺ complex is quite similar to that

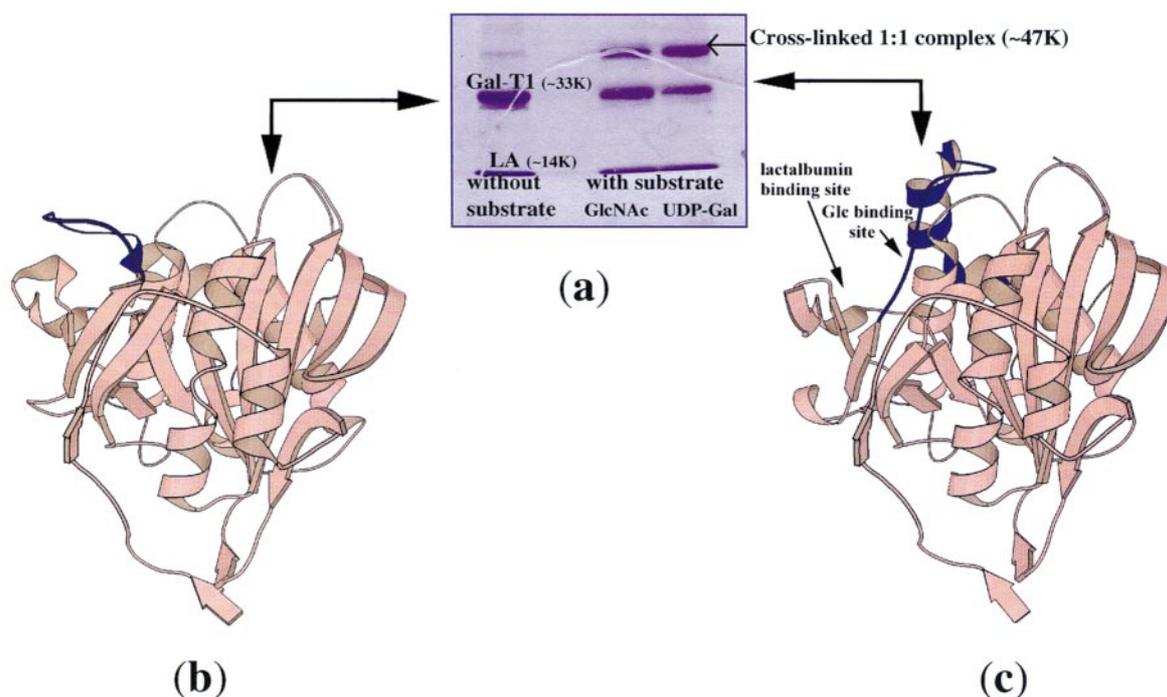


Figure 4. (a) SDS-PAGE analysis of the cross-linked catalytic domain of bovine β 4Gal-T1 and mouse LA. The cross-linking experiments were carried out as described by Brew *et al.*¹⁴ Left lane: cross-linking done in the absence of β 4Gal-T1 substrates. No higher molecular mass band corresponding to a 1:1 β 4Gal-T1 and LA complex is observed. Right two lanes: the cross-linking experiments were carried out in the presence of GlcNAc, and UDP-Gal and MnCl₂, respectively. Under these experimental conditions, β 4Gal-T1 and LA interact and can be cross-linked, producing a higher molecular mass band corresponding to a 1:1 complex. (b) The crystal structure of β 4Gal-T1 by itself¹⁰ (conformation I). The flexible region, residues 345 to 365 (blue), is folded in such a way that it buries the LA and sugar acceptor-binding sites on β 4Gal-T1. (c) The structure of β 4Gal-T1 in the LS crystals. The fold of the flexible region (blue) is quite different compared to conformation I. LA and Glc-binding sites on β 4Gal-T1 are now exposed.

found in the LS·Glc and LS·GlcNAc complexes, these changes can be induced either by the donor or by the acceptor substrate molecule.

Binding of the sugar acceptor GlcNAc and Glc molecules to LS

Since the crystal structure of free β 4Gal-T1,¹⁰ conformation I, is known only with the UDP complex, the sugar acceptor-binding site could not be inferred clearly from the crystal structure. Therefore, in order to understand the modulation mechanism by LA it was essential to determine the LS synthase crystal structure with Glc and with GlcNAc. In the crystal structure of the LS·GlcNAc complex, only one hydrogen bond was found between GlcNAc and LA; the N^{δ1} atom of His32 of LA forms a hydrogen bond with the O-1 hydroxyl group of the GlcNAc moiety (Figure 5(a)). All the other hydrogen bonds were between the hydroxyl groups of GlcNAc and the β 4Gal-T1 in the regions that comprise residues Lys279 to Tyr289, Trp314 to Asp319 and Arg359 to Ile363. Arg359, Phe360, and Ile363, located in the α 6 helix of β 4Gal-T1, form a hydrophobic pocket that accommodates and fully encloses the N-2 acetyl group of GlcNAc (Figure 5(a)). The N-2 nitrogen atom is hydrogen-bonded to the Asp319 side-chain carboxylate oxygen atom; its carbonyl oxygen atom to the N^ε nitrogen atom of the Arg359 side-chain, while the methyl group is surrounded by Phe360 and Ile363.

There are two LS molecules in the asymmetric unit of the LS·Glc complex, each containing a Glc molecule. From the electron density maps, Glc in one of the LS molecules could be identified clearly as being in the β -conformation (β -Glc) (Figure 5(b)). Since the conformation of the other Glc could not be identified unequivocally, it is denoted as α/β -Glc. The O-1 and O-2 hydroxyl groups of both Glc molecules in LS·Glc form a hydrogen bond with the N^{δ1} nitrogen atom of His32 of LA and the carboxylate group of Asp319 of β 4Gal-T1, respectively. In the interaction with Arg359, however, the two Glc molecules differ. The O-1 hydroxyl group of β -Glc is hydrogen-bonded to the N^η nitrogen atom of Arg359 (Figure 5(b)), whereas the O-1 hydroxyl group of α/β -Glc is hydrogen-bonded to Arg359 through a water molecule (Figure 5(c)). The hydrogen-bonding interactions of the remaining hydroxyl groups of the Glc molecule with β 4Gal-T1 are similar to those observed for GlcNAc.

UDP-Gal and Mn²⁺ binding to the LS molecule

In the studies reported by Gatinel *et al.*, although the free β 4Gal-T1 crystals were soaked with UDP-Gal and MnCl₂, only the UDP molecule could be located in the crystal structure¹⁰ and not Mn²⁺. Since the presence of metal ion is an absolute requirement for UDP-Gal binding as well as for catalysis,¹⁷ we determined two crystal structures of LS, both crystals grown in the presence of 17 mM UDP-Gal and MnCl₂ (Table 1). In one of the struc-

tures, we could locate only UDP (LS·UDP) but in the other both UDP and Mn²⁺ could be seen (LS·UDP·Mn²⁺, Figure 6(a)). The detection of UDP and not UDP-Gal in both the crystal structures could be due either to the hydrolysis of UDP-Gal or because Gal moiety is disordered in the crystals. In both the crystal structures, the nucleoside binding to β 4Gal-T1 is similar, but the interactions of phosphate groups are quite different. The orientation of the diphosphate group and its interactions with β 4Gal-T1 in LS·UDP where Mn²⁺ was not detected were, to some extent similar to those in β 4Gal-T1 bound with UDP in conformation I.¹⁰ However, in the structure in which Mn²⁺ was present (LS·UDP·Mn²⁺), the diphosphate group interactions with β 4Gal-T1 are quite different. The anionic oxygen atoms of the diphosphate group of UDP interact with β 4Gal-T1 through the Mn²⁺ and through a hydrogen bond with Trp314 (Figure 6(a)).

In the LS·UDP·Mn²⁺ structure, the Mn²⁺ is seen to coordinate with five atoms, two of them are the anionic oxygen atoms of phosphate groups of UDP and the other three are the carboxylate oxygen atom of Asp254, the N^{ε2} nitrogen atom of H347, and the S^δ sulfur atom of Met344 (Figure 6(a)). During refinement, the coordination distances were not constrained, and they varied from 2.2 to 2.9 Å. The involvement of the S^δ atom of Met residues in proteins in metal binding is not so common.¹⁸ The distance between the S^δ atom and Mn²⁺ is 2.9 Å, still within the observed coordination distance for the Mn²⁺. Among the β 1,4-Gal-T (galactosyltransferase) family members, His347 is highly conserved whereas Met344 is present only in β 4Gal-T1, -T2, -T3 and -T4.⁵ It has been thought that the sequence D²⁵²VD²⁵⁴ is involved in Mn²⁺ binding,¹⁹ and that this sequence is conserved in all the β 4Gal-T1 family members.¹⁹ However, the present crystal structure reveals that only Asp254 is involved in Mn²⁺ binding. It is possible that Asp252 is involved in galactose binding. Although we did not observe the donor galactose sugar in the crystal structure of LS with the UDP·Mn²⁺, its binding site can be visualized easily to be deep inside the catalytic pocket from the following fact: of the three β -phosphate anionic oxygen atoms, one is coordinated with the Mn²⁺, the second is hydrogen-bonded to Trp314, and the third oxygen atom pointing towards the interior of the pocket is likely to be galactose-bonded. In the pocket, the hydroxyl groups of galactose are expected to form extensive hydrogen bonding with the side-chain carboxylate oxygen atoms of Asp252, Glu317, and Asp318. Extension of the present work identifies galactose in the deep pocket in the crystals of the β 4Gal-T1 complex with UDP-Gal·Mn²⁺ (unpublished results).

Previous biochemical studies have shown that there are two metal-binding sites in β 4Gal-T1.^{17,20} The primary one is a high-affinity site that binds transition metal ions, particularly Mn²⁺, with K_D values in the μ M range. Ca²⁺ does not bind to this

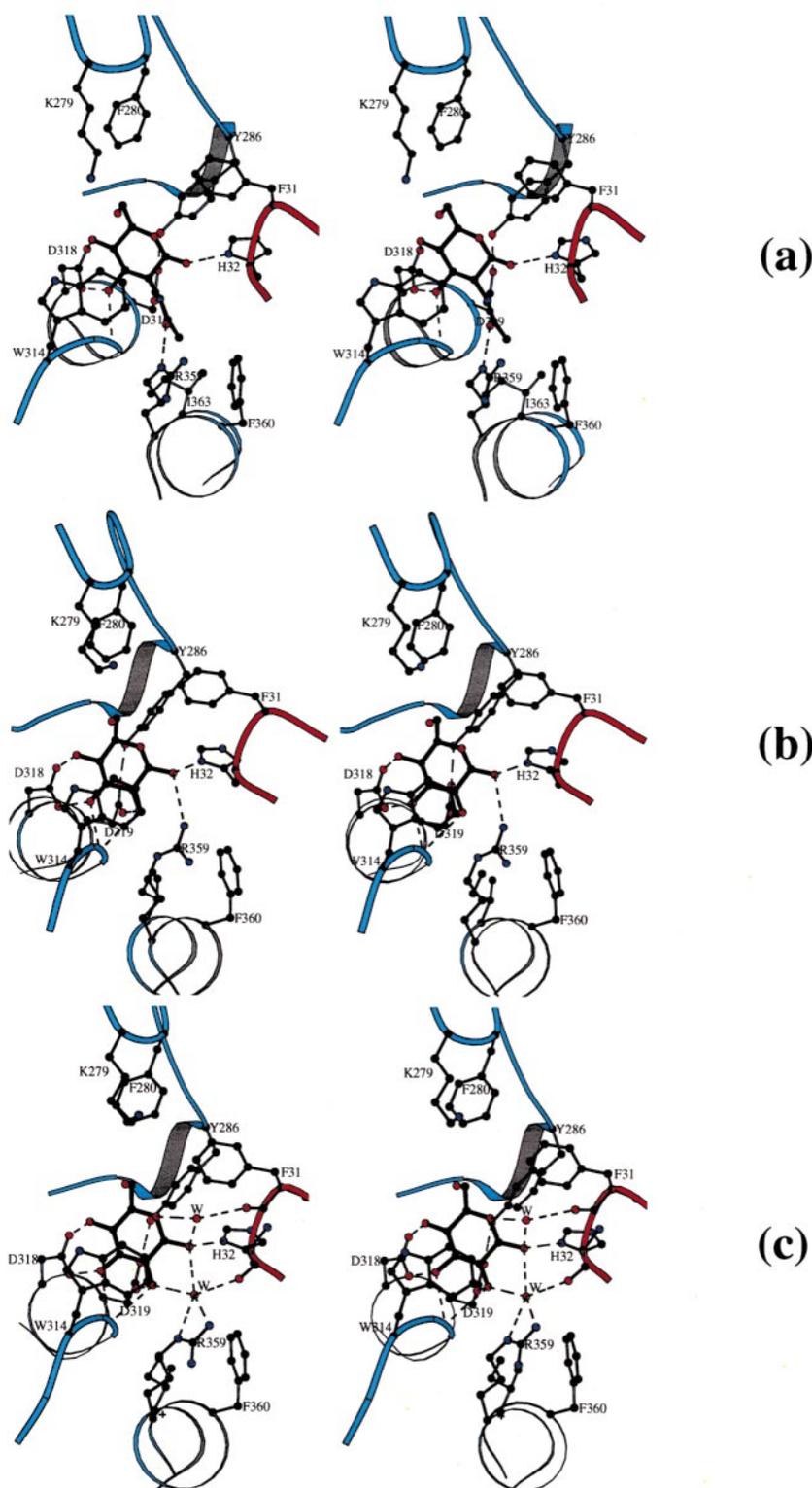


Figure 5. Molecular interactions involving the acceptor sugars (a) GlcNAc and (b) and (c) Glc in LS. The acceptors are shown in thick black color. (a) The *N*-acetyl group is being recognized through the residues Arg359, Phe360, and Ile363. These residues are in the α -helix conformation and form a hydrophobic pocket with the arginine guanidino group at the base of the pocket. The arginine side-chain NH atom forms a hydrogen bond with the carbonyl oxygen atom, while the methyl group is found inside the hydrophobic pocket. (b) In the presence of β -Glc, the O-1 hydroxyl group is exposed to the solvent, and Arg359 reaches out to form a hydrogen bond with it. Because of this, the hydrophobic pocket created by Phe360 and Ile363 gets closed. The absence of the *N*-acetyl group in Glc is compensated for by the formation of an additional hydrogen bond between the 2-hydroxyl atom to the backbone 314 residue. (c) With α/β -Glc, the O-1 hydroxyl group is partially buried and, hence, Arg359 cannot form a hydrogen bond with the O-1 hydroxyl group. Instead, the O-1 and O-2 hydroxyl groups of α/β -Glc interact with Arg359 through water molecules. Again, the particular side-chain conformation of arginine and the presence of the water molecule compensate for the lack of *N*-2 acetyl group in Glc.

site. The second metal-binding site is the low-affinity site with a K_D in the mM range, to which both Mn^{2+} and Ca^{2+} can bind.¹⁶ The primary binding site has been suggested to be important for the binding of UDP-Gal as well as for catalysis. The second metal-binding site, when occupied,

enhances the efficiency of catalysis and strengthens the acceptor binding.¹⁷

Even though we used 17 mM $MnCl_2$ for the crystallization of the LS·UDP· Mn^{2+} complex, we were unable to identify the second metal ion. However, a second metal ion-binding site is possibly located in the loop region between residues 311 to

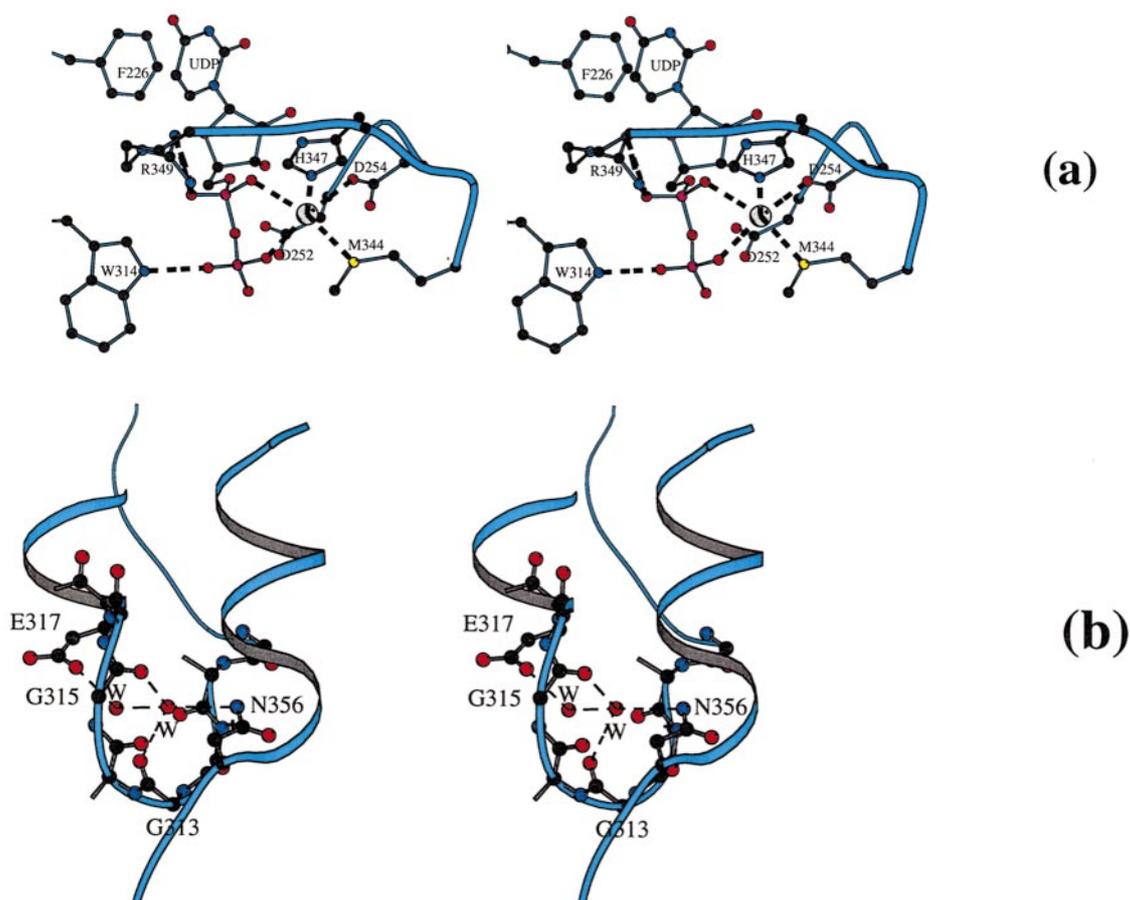


Figure 6. (a) Stereo view of UDP-Gal·Mn²⁺ binding to LS. The Mn²⁺ bridges UDP with β4Gal-T1 through five coordinations, of which two are to phosphate anionic oxygen atoms and three to the β4Gal-T1 molecule, Asp254, Met344 and His347. An unusual coordination is seen between Mn²⁺ and the side-chain S^δ atom of Met344 (see the text). (b) Stereo view of the possible second metal ion-binding site on β4Gal-T1. In all the crystal structures determined in the present study, a water molecule was found at the center of the loop between residues 311 and 317. This water molecule interacts with two backbone carbonyl oxygen atoms, another water molecule and with the side-chain amino group of Asn356. In the crystal structure of free β4Gal-T1¹⁰ this water molecule is absent, due to the different backbone conformation of the residues 313 and 314. The water molecule could play an important role in determining the side-chain orientation of Trp314.

317 (Figure 6(b)). This loop is sandwiched between the side-chains of Glu317 and Asn356. The loop in the conformation I of β4Gal-T1 has a different conformation for the residues Gly313 and Trp314 than does the conformation II and no water molecule is found at the center of the loop. Instead, in conformation I a water molecule was found close to the center, but outside the loop, interacting with the peptide backbone of the loop.¹⁰ Comparison of these two structures shows that residues Gly313 and Trp314 have undergone torsional changes, in such a way that their peptide backbones have flipped inside-out. Such conformational transition could create enough space at the center of the loop to accommodate a water molecule. It is possible that a metal ion can substitute for the water molecule and bring about the observed conformational change. This change alters the orientation of the

Trp314 side-chain from the exterior to the interior of the catalytic pocket. The side-chain orientation of Trp314 is important for UDP-Gal binding and for that of the sugar acceptor. This view is supported by the observation that β4Gal-T1 binds to a GlcNAc-agarose column better in the presence of Mn²⁺ than in its absence.¹⁷

It is possible to propose a model, based on the different crystal structures, for the binding of sugar acceptor and UDP-Gal·Mn²⁺ to β4Gal-T1 molecule (Figure 7). In this model, the sugar donor, UDP-Gal, and the sugar acceptor, Glc/GlcNAc, line up in the catalytic pocket where the O-4 hydroxyl group of the sugar acceptor, Glc/GlcNAc, is in the vicinity, 4-5 Å away from the C-1 atom of the donor Gal sugar of UDP-Gal. A very similar arrangement of the acceptor and donor sugars has been predicted on the basis of β4Gal-T1 inhibitor

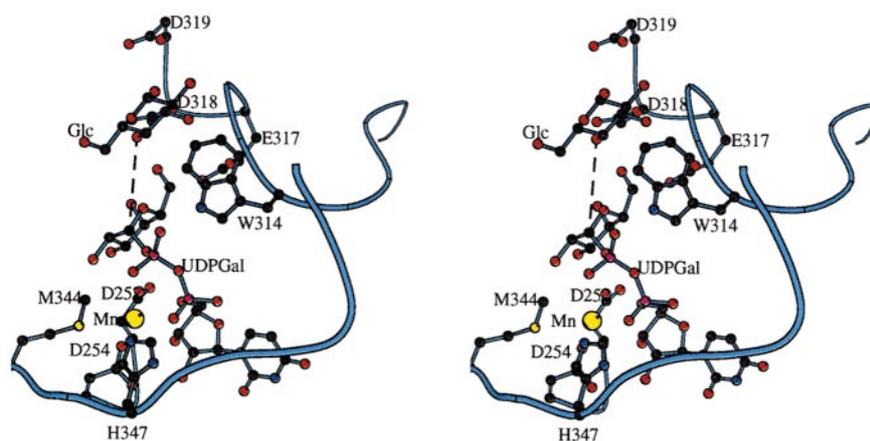


Figure 7. Stereo diagram of the model of the assembly of UDP-Gal, Mn^{2+} and Glc in the catalytic pocket of $\beta 4\text{Gal-T1}$ in the LS complexes. The interactions shown are based on the structures of individual substrate complexes. The O-4 hydroxyl group of the acceptor is within 4.5 Å of the C-1 atom of the Gal moiety in UDP-Gal.

studies.²¹ A similar acceptor-donor lineup in the catalytic pocket has been proposed for GlcNAc-T.²²

Relevance of the two conformational states for the enzyme kinetic pathway

The detailed enzyme kinetics studies^{23,24} on $\beta 4\text{Gal-T1}$ and LS have suggested that during catalysis the $\beta 4\text{Gal-T1} \cdot \text{Mn}^{2+}$ complex is formed first, to which UDP-Gal and sugar acceptor bind sequentially. In the presence of LA, LA and Glc molecules bind synergistically.²³ After catalysis, the product lactose and LA molecules dissociate, followed by UDP and Mn^{2+} . On the basis of the results of the present and the previous crystal structure¹⁰ of $\beta 4\text{Gal-T1}$ in conformation I, we propose a possible structural basis that accounts for the enzyme kinetic mechanism.

Binding of free UDP-Gal and Mn^{2+} to $\beta 4\text{Gal-T1}$ is possible only if $\beta 4\text{Gal-T1}$ is in conformation I, as in the free $\beta 4\text{Gal-T1}$ crystal structure¹⁰ (Figure 8(a)). This is supported by the fact that, for the elucidation of its structure, Gastinel *et al.* soaked free $\beta 4\text{Gal-T1}$ crystals with UDP-Gal and MnCl_2 to allow the diffusion of UDP-Gal into the molecule that formed a $\beta 4\text{Gal-T1} \cdot \text{UDP}$ complex. In conformation II, the UDP-Gal-binding site is partially blocked by the flexible region residues 345 to 365, preventing its binding to $\beta 4\text{Gal-T1}$. In solution, during the binding of UDP-Gal as a first step, the Trp314 of $\beta 4\text{Gal-T1}$ side-chain has to rotate from outside of the catalytic pocket, the orientation observed in conformation I, into the interior and lock UDP-Gal by forming a hydrogen bond with the β -phosphate oxygen atom (Figure 8(b)). The rotary motion can be achieved by torsional adjustment involving the highly conserved Gly313 residue as well as Trp314. Since in conformation II Trp314 side-chain partially blocks the UDP-Gal-binding site, UDP-Gal has to bind to $\beta 4\text{Gal-T1}$ prior to this conformational change. Next, the conformation and fold of the flexible region of $\beta 4\text{Gal-T1}$

(residues 345 to 365) changes from conformation I to II. This flexible region in conformation I had buried the LA and the acceptor-binding sites. When it adopts conformation II, it exposes the LA and the acceptor-binding sites and it covers the bound UDP-Gal (Figure 8(c)). Also, this conformational change brings His347 close to the Mn^{2+} so that it forms a coordination bond. This coordination is important to bind the Mn^{2+} ion to $\beta 4\text{Gal-T1}$ and to hold the flexible region in conformation II. This conformational change could be induced by the basic residues present in the region R³⁴⁹DKKN³⁵³, which are in the vicinity of the anionic phosphate groups of the bound UDP-Gal molecule. At this stage, both LA and acceptor sites are available for binding. Although, before LA molecule binds to $\beta 4\text{Gal-T1}$ in conformation II, a Glc molecule could sterically occupy its monosaccharide-binding site. It cannot by itself bind in the GlcNAc-binding site, since $\beta 4\text{Gal-T1}$ has low affinity for Glc without LA (see below). If the LA molecule binds first to $\beta 4\text{Gal-T1}$ in conformation II (which is possible), it would hinder the binding of the Glc molecule. Therefore, it is essential that both Glc and LA molecules bind to $\beta 4\text{Gal-T1}$ together in conformation II. In the final step, after Glc and LA-binding, $\beta 4\text{Gal-T1}$ undergoes further conformational changes on the Arg359 side-chain to maximize its interactions with the Glc molecule (see below). After catalysis, LA and the reaction product, lactose, leave. This is followed by the reversal of the conformation of the flexible region to conformation I in order to release the UDP and Mn^{2+} .

It has been shown that in the LS enzymatic activity, GlcNAc can be an acceptor only at lower concentrations (lower than its K_m value); at higher concentrations, it acts as a non-competitive inhibitor.²³ The structural mechanism proposed would explain this characteristic property also. At higher concentrations, GlcNAc by itself can induce the conformational changes from conformation I to II, prior to UDP-Gal binding to $\beta 4\text{Gal-T1}$. Steric

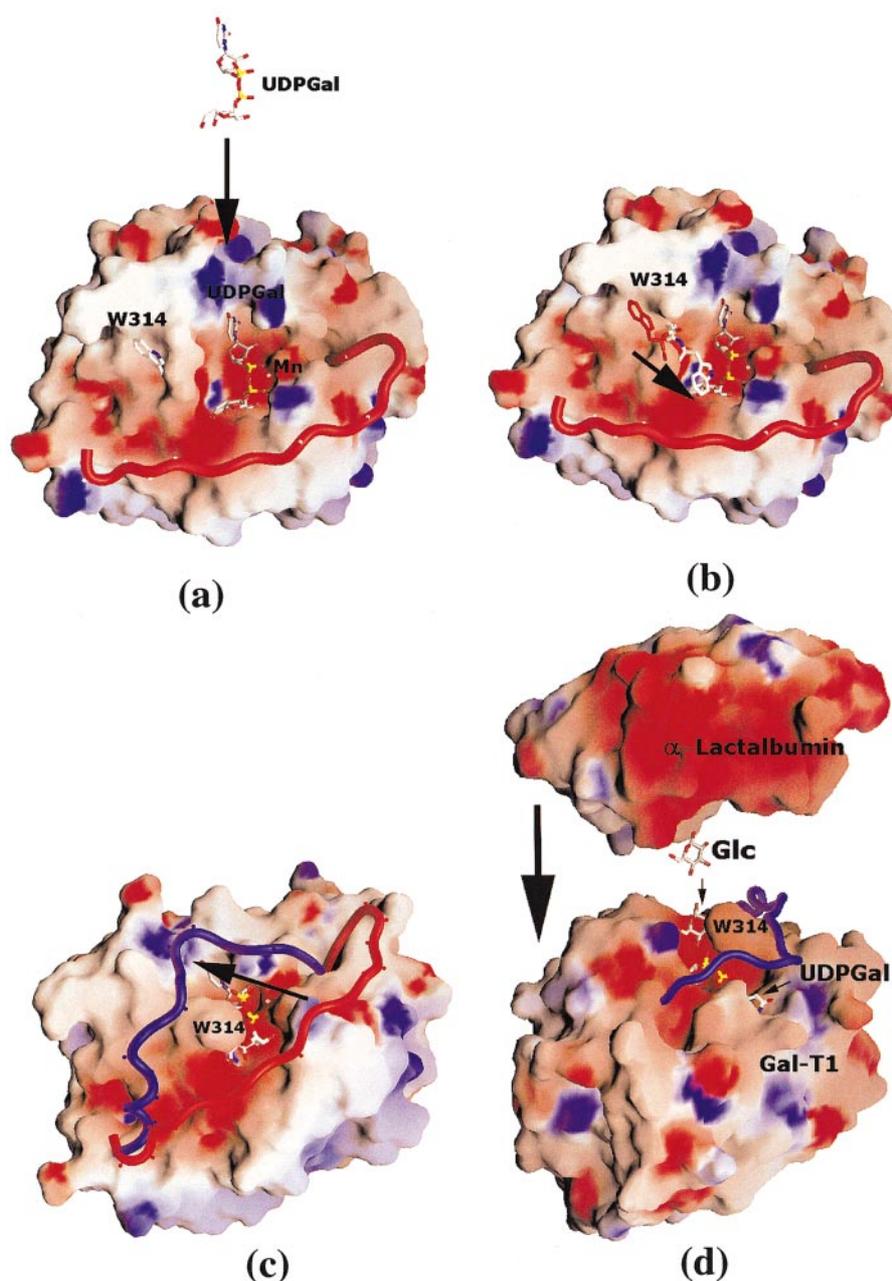


Figure 8. The intermediate stages of the catalytic activity are illustrated using the surface electrostatic potential of protein molecules. Electrostatic potential is mapped onto the molecular surface from -10kT (red) to $+10\text{kT}$ (blue). Based on the crystal structure of LS with the different substrates, conformation I of $\beta 4\text{Gal-T1}$,¹⁰ and the kinetic mechanism $\beta 4\text{Gal-T1}$ ²² established the plausible sequential binding of substrates to $\beta 4\text{Gal-T1}$ and the expected conformational transitions are shown. (a) The UDP-Gal molecule and Mn^{2+} bind to $\beta 4\text{Gal-T1}$ in conformation I. In this conformation its catalytic pocket is wide open for the entry of UDP-Gal and Mn^{2+} . (b) The Trp314 side-chain changes its orientation (red), from pointing to the outside (conformation I) to one located inside (white), as in conformation II. In this state, the secondary N of the indole group of Trp314 hydrogen bonds to the β -phosphate anionic oxygen atom, thereby properly orienting the phosphate groups and allowing its conformation to change. (c) The conformational change of the flexible region, residues 345 to 365, from conformation I (red) to active state conformation II (blue). This alters the secondary structure of region 358 to 365 from a loop to α -helix. This also creates the sugar acceptor pocket and exposes the LA-binding site. (d) To form the LS complex with the sugar acceptor and to achieve the active state, LA has to bind along with Glc to the UDP-Gal and Mn^{2+} bound $\beta 4\text{Gal-T1}$ (see the text). After step (d), LS is expected to transfer Gal from UDP-Gal to Glc, to synthesize lactose.

conditions preclude the UDP-Gal molecule from binding to the $\beta 4\text{Gal-T1}$ -GlcNAc complex, which is in conformation II. This complex is catalytically

a “dead-end” state of the molecule, indicative of the non-competitive inhibition of the enzyme at high concentrations of GlcNAc. Similarly, the

observed inhibition of LS catalysis at higher concentrations of LA can be explained as follows. The binding of the LA molecule to the active-state β 4Gal-T1 molecule in conformation II without a Glc molecule, which forms the β 4Gal-T1·UDP-Gal·Mn²⁺·LA complex, is also a catalytically dead-end molecule, as Glc cannot bind to this complex for steric reasons.

Molecular basis of the modulation of sugar acceptor specificity of β 4Gal-T1 by LA

The modulation of the β 4Gal-T1 catalytic activity by LA in the LS complex consists of two components: (1) the sugar acceptor preference of β 4Gal-T1 for Glc is enhanced 1000-fold; and (2) the preference for a monosaccharide acceptor over an oligosaccharide. The latter can be explained on the basis of the molecular interactions between β 4Gal-T1 and LA molecules. Since LA binds at the extended region of the monosaccharide-binding site in β 4Gal-T1, GlcNAc containing extended sugars, such as GlcNAc(β 1-4)GlcNAc are not acceptor substrates for the LS molecule. Thus, LA and extended sugars compete for the same site on the β 4Gal-T1 molecule and become competitive inhibitors.^{24,25}

The preference for Glc in LS can be understood from the crystal structures of LS. Even in the absence of the acceptor N-2 acetyl-glucosamine (GlcNAc), the LA·Gal-T1·UDP-Gal·Mn²⁺ crystal structure shows a hydrophobic pocket at the sugar acceptor-binding site of β 4Gal-T1, which would facilitate the binding of N-2 acetyl group of GlcNAc (Figure 9(a)). In LS·GlcNAc this hydrophobic pocket is occupied by the N-2-acetyl group of GlcNAc (Figure 9(b)). If the glucose molecule were to bind to the LS in this form, it can be expected to bind only weakly in the absence of the N-2 acetyl group to fill the hydrophobic pocket. But in the crystal structure of the LS·Glc complex, this hydrophobic pocket is absent due to the reorientation of the side-chain of Arg359 (Figure 9(c) and (d)). Thus, it seems that the presence or absence of the N-2 acetyl group on the sugar acceptor is being recognized solely by the β 4Gal-T1 molecule. The only direct interaction between the LA and Glc molecules in the LS·Glc complex is a hydrogen bond between His32 of LA and the O-1 hydroxyl group of Glc (Figure 9(c) and (d)). This hydrogen bond, however, is not specific for Glc only, since it is present with GlcNAc (Figure 9(b)). Thus, the role of LA seems to be to block the extended sugar acceptor-binding pocket, while leaving only a monosaccharide-binding site open and to hold the monosaccharide by a hydrogen bond. It appears that while Glc is being held in the acceptor-binding pocket by LA, β 4Gal-T1 modulates itself by adjusting the orientation of Arg359 side-chain to maximize its interactions with the Glc molecule. This way, the acceptor specificity of β 4Gal-T1 is broadened to include Glc along with GlcNAc as an acceptor in the LS complex.

β 4Gal-T1 exists as a family of transferases with seven members, β 4Gal-T1 to -T7.^{5,6} Although only β 4Gal-T1 is expressed in the mammary gland during lactation, it has been shown that LA does not effectively modulate the acceptor specificity of several other family members; with β 4Gal-T4 it enhances the transfer of Gal to GlcNAc more than to Glc.⁷ Since the LA-binding residues are present in most of these family members, attempts are being made, based on the present β 4Gal-T1 structure, to explain LA-binding to other β 4Gal-T family members by sequence homology model building. Such studies are expected to give some insight into the evolutionary relationship that exists among the β 4Gal-T family members. It is interesting to note that β 4Gal-T6, which transfers Gal from UDP-Gal to glucosylceramide, has N-acetyl binding residues Arg359, Phe360, and Ile363. It is possible, therefore, that the N-acetyl binding pocket is present on β 4Gal-T6. In order to transfer Gal to Glc in the β 4Gal-T6 reaction, the ceramide group of the acceptor may play a role similar to that of LA in LS.

Materials and Methods

β 4Gal-T1 and mouse LA expression and purification

The catalytic domain of recombinant bovine β 4Gal-T1 from residues, 130 to 402, d129 β 4Gal-T1 (~33 kDa),²⁶ and mouse recombinant LA (~14 kDa) were expressed in *E. coli* as inclusion bodies. We used an S-sulfonation protocol for folding the protein from the inclusion bodies; a method previously used successfully for folding recombinant phospholipase A2.²⁷ Inclusion bodies (100 mg) were dissolved in 10 ml of 5 M guanidine hydrochloride (Gu-HCl) with 0.3 M sodium sulfite at room temperature. To sulfonate all the free thiol groups in the protein molecule, 1 ml of 50 mM S-sulfonating agent, 2-nitro-5-(sulfothio)-benzoate (NTSB), was added to this solution and stirred vigorously. Completion of sulfonation was judged by the color change of the solution from red to pale yellow. The protein solution was then diluted tenfold with water to precipitate the sulfonated protein. The protein precipitate was collected by centrifugation at 10,000 g, washed three times by re-suspending in water, followed by centrifugation to remove any remaining sulfonating agent. The sulfonated protein was re-dissolved in 5 M Gu-HCl to a protein concentration of 1 mg/ml, which has an absorption of 1.9 to 2.0 at 275 nm. The protein solution was diluted tenfold, in 10 ml portions, in a folding solution to give a final concentration of 100 μ g/ml, 0.5 M Gu-HCl, 50 mM Tris-HCl (pH 8.0 at 4°C), 5 mM EDTA, 4 mM cysteamine, 2 mM cystamine. The protein was allowed to fold for 48 hours at 4°C, then dialyzed against three changes of 4 l of 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, 4 mM cysteamine, 2 mM cystamine at 4°C to remove Gu-HCl. The protein that precipitated during dialysis was removed by centrifugation and the supernatant was concentrated. Typically, when 100 mg of sulfonated wild-type d129- β 4Gal-T1 is folded in a 1 l folding solution, it yields 3 to 5 mg of active, soluble, and pure protein. The folded active

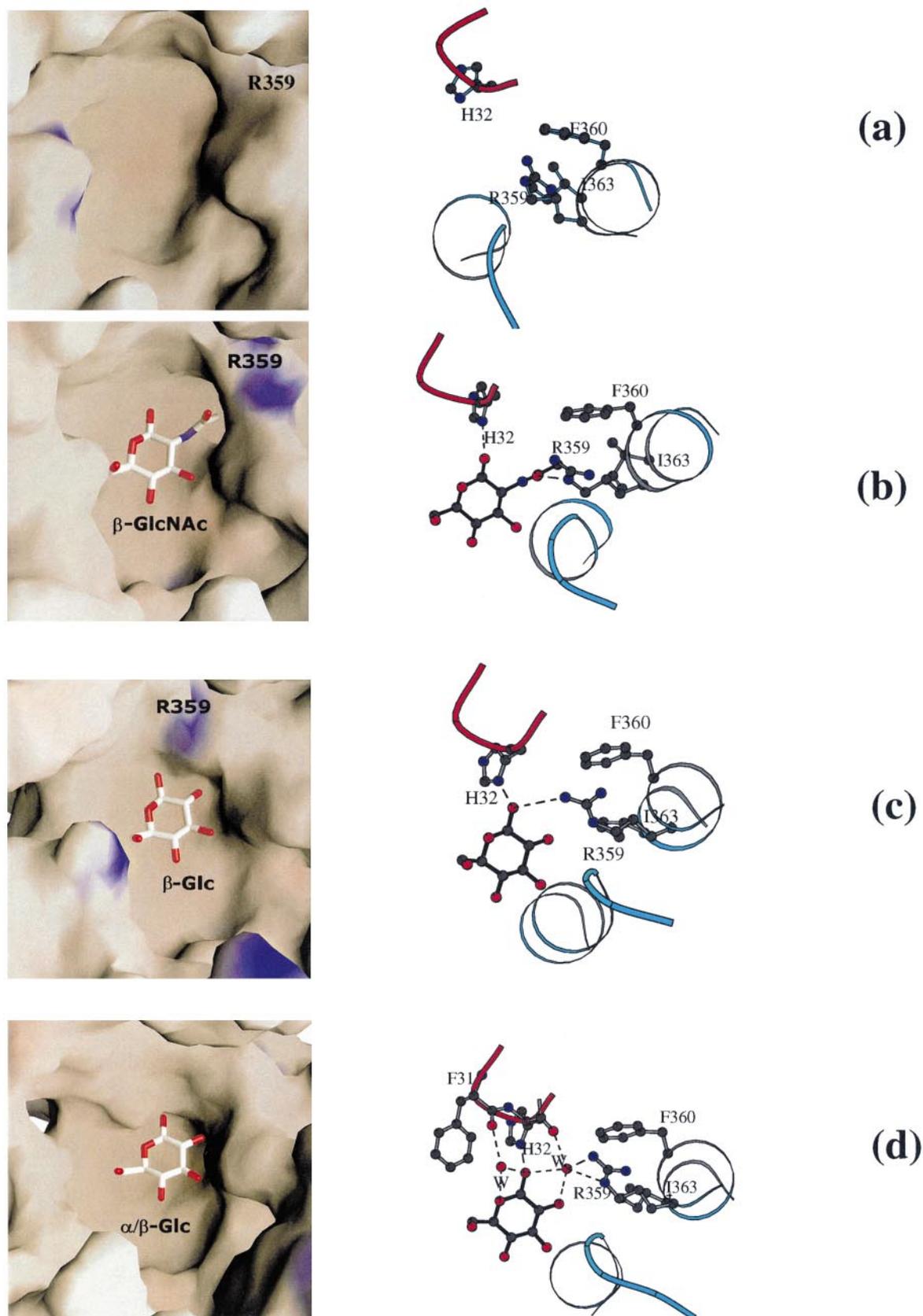


Figure 9. Molecular surface diagram and interactions showing differences in the acceptor binding pocket at the second exo-cyclic position of GlcNAc and Glc. (a) In the LS·UDP·Mn²⁺ complex, which does not contain any sugar acceptor, the *N*-acetyl group binding pocket is still present. (b) In the LS-GlcNAc complex, when the GlcNAc molecule is present the hydrophobic pocket is completely occupied by the *N*-acetyl group of GlcNAc. In the absence of the *N*-acetyl group, this pocket will remain empty if Glc were to bind in this site. (c) and (d) In the LS-Glc complex, Glc binds in this region where the side-chain conformation of Arg359 adopts an orientation that closes the hydrophobic pocket.

β 4Gal-T1 was further purified by affinity chromatography using an LA-agarose column.

Unlike β 4Gal-T1, the S-sulfonated inclusion bodies of LA remained soluble and were used as such for folding. Guanidine hydrochloride was required neither for solubilization nor for folding of sulfonated protein. The concentration of the dialyzed sulfonated LA solution was adjusted to about 1 mg/ml by monitoring its absorption at 275 nm (2.4 to 2.6). Calcium chloride has been shown²⁸ to be required for the folding of the recombinant LA. Typically, an equal amount of a 2 \times folding solution containing 100 mM sodium borate (pH 8.5), 5 mM calcium chloride, 8 mM reduced glutathione, 4 mM oxidized glutathione was added slowly to the protein solution. The resulting solution was left to fold at room temperature for 48 hours and then dialyzed against 1 mM calcium chloride to remove the buffer and the oxido-shuffling agents. The dialyzed protein solution was centrifuged at 15,000 g to remove any traces of precipitates and then concentrated. The active protein was isolated from the partially folded inactive protein by fractional precipitation at 40-70% saturated ammonium sulfate. The protein precipitate was dissolved in water and dialyzed extensively against 1 mM calcium chloride, and the resulting protein solution was directly used for crystallization and other studies. Using this method, 60 to 70% of the starting S-sulfonated LA material was obtained in an active form.

Crystallization and data collection

Crystals were grown at room temperature by the hanging drop methods, using 20 mg ml⁻¹ of β 4Gal-T1 and 10 mg ml⁻¹ of LA in the presence of a substrate with the precipitant containing 100 mM NaCl, 100 mM sodium citrate (pH 5.6) and 12.5% (w/v) PEG 4000. The crystals of the complex could be obtained only in the presence of substrates. The concentration of substrates used during the crystallization were 10 mM GlcNAc, 100 mM Glc and UDP-Gal, and 17 mM MnCl₂. Complete three-dimensional X-ray diffraction data were collected at beam line X9B, NSLS, BNL, using a Quantum-4 ccd detector. The frames were processed using DENZO.²⁹ For solving the crystal structure of LS, we used Se-met containing β 4Gal-T1 to grow LS crystals. The MAD data were collected at three wavelengths, 0.96, 0.98 and 0.99 Å, up to 2.5 Å resolution. The native data on the other substrate-bound LS crystals were collected using a wavelength of 0.98 Å. The data collection statistics are given in Table 1.

Structure solution and refinement

The structure was solved using MAD methods. Of the 22 Se atoms, 14 were located using SOLVE³⁰ at 2.8 Å resolution. Solvent averaging of Fourier maps based on these Se atoms by DM in CCP4³¹ yielded interpretable electron density maps. Although no Se atom is present in the LA molecule, the experimental electron density maps at the 1 σ level clearly showed both LA and β 4Gal-T1 molecules. Based on the experimental electron density maps, the individual crystal structures of the LA and the β 4Gal-T1 molecules were used as models to build the LS structure. The crystal structure of rat LA (unpublished results) was fitted into the electron density map of LA without any change in its conformation. Although the β 4Gal-T1 structure¹⁰ could be fitted into the experimental

electron density map, a large difference in conformation and fold for residues 313 to 314 and 345 to 365 was observed. Therefore, these regions alone were built independently using the program O.³² The two LS molecules in the asymmetric unit are related by a pseudo 2-fold symmetry. During refinement, the two LS molecules are restrained by this pseudo 2-fold symmetry. All the refinements were carried out initially by XPLOR3.85, followed by CNS1.0.³³ The coordination distances for Mn²⁺ were not restrained, while the Ca²⁺ coordination distances in LA were restrained. The final refinement statistics are given in Table 1. All the Figures were drawn using MOLSCRIPT³⁴ and the electrostatic surface diagrams were generated using GRASP.³⁵

Protein Data Bank accession numbers

The coordinates of LS·GlcNAc, LS·Glc, LS·UDP·Mn and Se-LS·UDP complexes have been deposited in the RCSB Brookhaven Protein Data Bank with accession numbers 1J92, 1J8W, 1J8X and 1J94.

Acknowledgments

We thank Dr Zbigniew Dauter for help with synchrotron data collection. We thank Drs David Davies, Alexander Wlodawer, Robert Hill, Jacob Maizel, Soma Kumar and Elizabeth Boeggeman for helpful discussions and comments during preparation of the manuscript. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organization imply endorsement by the US Government. This project has been funded, in part, with Federal funds from the NCI, NIH, under contract no. NO1-CO-56000.

©US Government

References

1. Brew, K., Vanaman, T. C. & Hill, R. L. (1968). The role of alpha-lactalbumin and the A protein in lactose synthetase: a unique mechanism for the control of a biological reaction. *Proc. Natl Acad. Sci. USA*, **59**, 491-497.
2. Brodbeck, U., Denton, W. L., Tanahashi, N. & Ebner, K. E. (1967). The isolation and identification of the B protein of lactose synthetase as alpha-lactalbumin. *J. Biol. Chem.* **242**, 1391-1397.
3. Klee, W. A. & Klee, C. B. (1970). The role of alpha-lactalbumin in lactose synthetase. *Biochem. Biophys. Res. Commun.* **39**, 833-834.
4. Qasba, P. K. & Kumar, S. (1997). Molecular divergence of lysozymes and alpha-lactalbumin. *Crit. Rev. Biochem. Mol. Biol.* **32**, 255-306.
5. Lo, N. W., Shaper, J. H., Pevsner, J. & Shaper, N. L. (1998). The expanding beta 4-galactosyltransferase gene family: messages from the databanks. *Glycobiology*, **8**, 517-526.
6. Amado, M., Almeida, R., Schwientek, T. & Clausen, H. (1998). Identification and characterization of large galactosyltransferase gene families: galactosyltransferases for all functions. *Biochim. Biophys. Acta*, **1473**, 35-53.
7. Schwiente, T., Almeida, R., Levery, S. B., Holmes, E. H., Bennett, E. & Clausen, H. (1998). Cloning of a

- novel member of the UDP-galactose:beta-N-acetylglucosamine beta1,4-galactosyltransferase family, beta4Gal-T4, involved in glycosphingolipid biosynthesis. *J. Biol. Chem.* **273**, 29331-29340.
8. Nomura, T., Takizawa, M., Aoki, J., Arai, H., Inoue, R. & Wakisaka, R. *et al.* (1998). Purification, cDNA cloning, and expression of UDP-Gal: glucosylceramide beta-1,4-galactosyltransferase from rat brain. *J. Biol. Chem.* **273**, 13570-13577.
 9. Acharya, K. R., Stuart, D. I., Walker, N. P., Lewis, M. & Phillips, D. C. (1989). Refined structure of baboon alpha-lactalbumin at 1.7 Å resolution. Comparison with C-type lysozyme. *J. Mol. Biol.* **208**, 99-127.
 10. Gastinel, L. N., Cambillau, C. & Bourne, Y. (1999). Crystal structures of the bovine beta4galactosyltransferase catalytic domain and its complex with uridine diphosphogalactose. *EMBO J.* **18**, 3546-3557.
 11. Clymer, D. J., Geren, R. & Ebner, K. E. (1976). Ultraviolet photoinactivation of galactosyltransferase. Protection by substrates. *Biochemistry*, **15**, 1093-1097.
 12. Malinowski, V. A., Tian, J., Grobler, J. A. & Brew, K. (1996). Functional site in alpha-lactalbumin encompasses a region corresponding to a subsite in lysozyme and parts of two adjacent flexible substructures. *Biochemistry*, **35**, 9710-9715.
 13. Harata, K. & Muraki, M. (1992). X-ray structural evidence for a local helix-loop transition in alpha-lactalbumin. *J. Biol. Chem.* **267**, 1419-1421.
 14. Pike, A. C., Brew, K. & Acharya, K. R. (1996). Crystal structures of guinea-pig, goat and bovine alpha-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure*, **4**, 691-703.
 15. Brew, K., Shaper, J. H., Olsen, K. W., Trayer, I. P. & Hill, R. L. (1975). Cross-linking of the components of lactose synthetase with dimethylpimelidate. *J. Biol. Chem.* **250**, 1434-1444.
 16. Geren, C. R., Magee, S. C. & Ebner, K. E. (1975). Circular dichroism changes in galactosyltransferase upon substrate binding. *Biochemistry*, **14**, 1461-1463.
 17. Powell, J. T. & Brew, K. (1976). Metal ion activation of galactosyltransferase. *J. Biol. Chem.* **251**, 3645-3652.
 18. Adman, E. T., Turley, S., Bramson, R., Petratos, K., Banner, D. & Tsernoglou, D. *et al.* (1989). 2.0-Å structure of the blue copper protein (cupredoxin) from *Alcaligenes faecalis* S-6. *J. Biol. Chem.* **264**, 87-99.
 19. Wiggins, C. A. & Munro, S. (1998). Activity of the yeast MNN1 alpha-1,3-mannosyltransferase requires a motif conserved in many other families of glycosyltransferases. *Proc. Natl Acad. Sci. USA*, **95**, 7945-7950.
 20. O'Keeffe, E. T., Hill, R. L. & Bell, J. E. (1980). Active site of bovine galactosyltransferase: kinetic and fluorescence studies. *Biochemistry*, **19**, 4954-4962.
 21. Chung, S. J., Takayama, S. & Wong, C. H. (1998). Acceptor substrate-based selective inhibition of galactosyltransferases. *Bioorg. Med. Chem. Letters*, **8**, 3359-3364.
 22. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. & Rini, J. M. (2000). X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J.* **19**, 5269-5280.
 23. Khatra, B. S., Herries, D. G. & Brew, K. (1974). Some kinetic properties of human-milk galactosyltransferase. *Eur. J. Biochem.* **44**, 537-560.
 24. Bell, J. E., Beyer, T. A. & Hill, R. (1976). The kinetic mechanism of bovine milk galactosyltransferase. The role of alpha-lactalbumin. *J. Biol. Chem.* **251**, 3003-3013.
 25. Powell, J. T. & Brew, K. (1976). A comparison of the interactions of galactosyltransferase with a glycoprotein substrate (ovalbumin) and with alpha-lactalbumin. *J. Biol. Chem.* **251**, 3653-3663.
 26. Boeggeman, E. E., Balaji, P. V., Sethi, N., Masibay, A. S. & Qasba, P. K. (1993). Expression of deletion constructs of bovine beta-1,4-galactosyltransferase in *Escherichia coli*: importance of Cys134 for its activity. *Protein Eng.* **6**, 779-785.
 27. Noel, J. P., Bingman, C. A., Deng, T. L., Dupureur, C. M., Hamilton, K. J. & Jiang, R. T. *et al.* (1991). Phospholipase A2 engineering. X-ray structural and functional evidence for the interaction of lysine-56 with substrates. *Biochemistry*, **30**, 11801-11811.
 28. Wang, M., Scott, W. A., Rao, K. R., Udey, J., Conner, G. E. & Brew, K. (1989). Recombinant bovine alpha-lactalbumin obtained by limited proteolysis of a fusion protein expressed at high levels in *Escherichia coli*. *J. Biol. Chem.* **264**, 21116-21121.
 29. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
 30. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849-861.
 31. Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760 (1994).
 32. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these methods. *Acta Crystallog. sect. A*, **47**, 110-119.
 33. Brunger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W. & Jiang, J. S. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905-921.
 34. Karulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
 35. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.

Edited by R. Huber

(Received 20 February 2001; received in revised form 8 May 2001; accepted 8 May 2001)